

PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/914,151

DATE: 09/13/2001

TIME: 17:37:14

Input Set : A:\ES.txt

Output Set: N:\CRF3\09132001\I914151.raw

3 <110> APPLICANT: Juridical Foundation, Japanese Foundation For Cancer Research
 5 <120> TITLE OF INVENTION: vector for gene therapy of malignant melanoma, with use of
 virus h

6 having MSH fused protein.

8 <130> FILE REFERENCE: H11-0241J2

C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/914,151

C--> 10 <141> CURRENT FILING DATE: 2001-08-24

10 <160> NUMBER OF SEQ ID NOS: 39

12 <170> SOFTWARE: PatentIn Ver. 2.0

14 <210> SEQ ID NO: 1

15 <211> LENGTH: 166

16 <212> TYPE: DNA

17 <213> ORGANISM: Artificial Sequence

19 <220> FEATURE:

20 <223> OTHER INFORMATION: DNA coding a part of adenovirus type 5 fiber, AS linker
 peptide an

21 d Y-MSH.

23 <220> FEATURE

24 <221> NAME: US

25 <222> LOCATION: (3)..(113) OK

27 <400> SEQUENCE: 1

28 (gg gaa ttc tcg agt tac act ttt tca tac att gcc caa gaa cca tca 47

29 OK Glu Phe Ser Ser Tyr Thr Phe Ser Tyr Ile Ala Gln Glu Pro Ser

30 1 5 10 15 95

31 gcc tcc gca tct gct tcc gcc cct gga tcc tac tcc atg gag cac ttc

32 Ala Ser Ala Ser Ala Ser Ala Pro Gly Ser Tyr Ser Met Glu His Phe.

W--> 33 20 25 30→

34 cgc tgg ggc aag ccg gtg taaagaatcg tttgtgttat gtttcaacgt 143

35 Arg Trp Gly Lys Pro Val

W--> 36 35

37 gtttattttt caattgaatt ccc 166

40 <210> SEQ ID NO: 2

41 <211> LENGTH: 126

42 <212> TYPE: DNA

43 <213> ORGANISM: Artificial Sequence

45 <220> FEATURE:

46 <223> OTHER INFORMATION: synthetic DNA No.924 used as template for PCR amplification
 of DNA

47 sequence No.1.

49 <400> SEQUENCE: 2

50 cggtgaaaca taacacaaac gattctttac accggcttgc cccagcggaa gtgctccatg 60

51 gaggtaggagc caggggagga agcagatgag gaggtgatg gttcttgggc aatgtatgaa 120

52 aaagtg 126

55 <210> SEQ ID NO: 3

56 <211> LENGTH: 39

57 <212> TYPE: DNA

58 <213> ORGANISM: Artificial Sequence

60 <220> FEATURE:

61 <223> OTHER INFORMATION: synthetic DNA No.933 used as sense primer for PCR
amplification of
62 DNA sequence No.1.

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/914,151

DATE: 09/13/2001

TIME: 17:37:14

Input Set : A:\ES.txt

Output Set: N:\CRF3\09132001\I914151.raw

```

64 <400> SEQUENCE: 3
65   gggaattctc gagttacact ttttcataca ttgcccaag           39
68 <210> SEQ ID NO: 4
69 <211> LENGTH: 49
70 <212> TYPE: DNA
71 <213> ORGANISM: Artificial Sequence
73 <220> FEATURE:
74 <223> OTHER INFORMATION: synthetic DNA No.934 used as antisense primer for PCR
amplification
75   of DNA sequence No.1.
77 <400> SEQUENCE: 4
78   gggaattcaa ttgaaaaata aacacgttga aacataacac aaacgattc           49
81 <210> SEQ ID NO: 5
82 <211> LENGTH: 76
83 <212> TYPE: DNA
84 <213> ORGANISM: Artificial Sequence
86 <220> FEATURE:
87 <223> OTHER INFORMATION: synthetic DNA No.1061 used as sense primer for PCR
amplification of
88   DNA coding Y"-MSH and adenovirus fiber poly A signal.
90 <400> SEQUENCE: 5
91   cgggatccta ctccatggag cacttccgct ggggcaagcc ggtgtaagtc gacaagaata           60
92   aagaatcggtt tgtgtt                                           76
94 <210> SEQ ID NO: 6
95 <211> LENGTH: 32
96 <212> TYPE: DNA
97 <213> ORGANISM: Artificial Sequence
99 <220> FEATURE:
100 <223> OTHER INFORMATION: synthetic DNA No.1092 used as antisense primer for PCR
amplificati
101   on of DNA coding Y"-MSH and adenovirus fiber poly A signal.
103 <400> SEQUENCE: 6
104   cggaattcat ggcgcatgt ttaatcagag gt           32
107 <210> SEQ ID NO: 7
108 <211> LENGTH: 1818
109 <212> TYPE: DNA
110 <213> ORGANISM: Artificial Sequence
112 <220> FEATURE:
113 <223> OTHER INFORMATION: DNA coding a modified fiber protein of pWE6.7R-F/asMSHa
115 <220> FEATURE:
116 <221> NAME/KEY: CDS
117 <222> LOCATION: (1)..(1815)
119 <400> SEQUENCE: 7
120   atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca           48
121   Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
122   1           5           10           15
124   tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc           96
125   Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
126   20           25           30
128   ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct           144
129   Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
130   35           40           45

```

RAW SEQUENCE LISTING

DATE: 09/13/2001

PATENT APPLICATION: US/09/914,151

TIME: 17:37:14

Input Set : A:\ES.txt

Output Set: N:\CRF3\09132001\I914151.raw

132	ttg	cgc	cta	tcc	gaa	cct	cta	gtt	acc	tcc	aat	ggc	atg	ctt	gcg	ctc	192
133	Leu	Arg	Leu	Ser	Glu	Pro	Leu	Val	Thr	Ser	Asn	Gly	Met	Leu	Ala	Leu	
134		50					55					60					
136	aaa	atg	ggc	aac	ggc	ctc	tct	ctg	gac	gag	gcc	ggc	aac	ctt	acc	tcc	240
137	Lys	Met	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Glu	Ala	Gly	Asn	Leu	Thr	Ser	
138		65				70					75				80		
140	caa	aat	gta	acc	act	gtg	agc	cca	cct	ctc	aaa	aaa	acc	aag	tca	aac	288
141	Gln	Asn	Val	Thr	Thr	Val	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Ser	Asn	
142					85					90					95		
144	ata	aac	ctg	gaa	ata	tct	gca	ccc	ctc	aca	gtt	acc	tca	gaa	gcc	cta	336
145	Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu	
146				100					105					110			
148	act	gtg	gct	gcc	gcc	gca	cct	cta	atg	gtc	gcg	ggc	aac	aca	ctc	acc	384
149	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr	
150			115					120					125				
152	atg	caa	tca	cag	gcc	ccg	cta	acc	gtg	cac	gac	tcc	aaa	ctt	agc	att	432
153	Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile	
154		130					135					140					
156	gcc	acc	caa	gga	ccc	ctc	aca	gtg	tca	gaa	gga	aag	cta	gcc	ctg	caa	480
157	Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln	
158		145				150					155				160		
160	aca	tca	ggc	ccc	ctc	acc	acc	acc	gat	agc	agt	acc	ctt	act	atc	act	528
161	Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr	
162				165					170					175			
164	gcc	tca	ccc	cct	cta	act	act	gcc	act	ggt	agc	ttg	ggc	att	gac	ttg	576
165	Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu	
166				180					185					190			
168	aaa	gag	ccc	att	tat	aca	caa	aat	gga	aaa	cta	gga	cta	aag	tac	ggg	624
169	Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly	
170			195					200					205				
172	gct	cct	ttg	cat	gta	aca	gac	gac	cta	aac	act	ttg	acc	gta	gca	act	672
173	Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr	
174			210				215						220				
176	ggt	cca	ggt	gtg	act	att	aat	aat	act	tcc	ttg	caa	act	aaa	gtt	act	720
177	Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr	
178		225				230					235				240		
180	gga	gcc	ttg	ggt	ttt	gat	tca	caa	ggc	aat	atg	caa	ctt	aat	gta	gca	768
181	Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala	
182				245					250					255			
184	gga	gga	cta	agg	att	gat	tct	caa	aac	aga	cgc	ctt	ata	ctt	gat	gtt	816
185	Gly	Gly	Leu	Arg	Ile	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val	
186			260						265					270			
188	agt	tat	ccg	ttt	gat	gct	caa	aac	caa	cta	aat	cta	aga	cta	gga	cag	864
189	Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln	
190			275					280						285			
192	ggc	cct	ctt	ttt	ata	aac	tca	gcc	cac	aac	ttg	gat	att	aac	tac	aac	912
193	Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn	
194			290				295					300					
196	aaa	ggc	ctt	tac	ttg	ttt	aca	gct	tca	aac	aat	tcc	aaa	aag	ctt	gag	960

RAW SEQUENCE LISTING

DATE: 09/13/2001

PATENT APPLICATION: US/09/914,151

TIME: 17:37:14

Input Set : A:\ES.txt

Output Set: N:\CRF3\09132001\I914151.raw

197	Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu	
198	305 310 315 320	
200	gtt aac cta agc act gcc aag ggg ttg atg ttt gac gct aca gcc ata	1008
201	Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile	
202	325 330 335	
204	gcc att aat gca gga gat ggg ctt gaa ttt ggt tca cct aat gca cca	1056
205	Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro	
206	340 345 350	
208	aac aca aat ccc ctc aaa aca aaa att ggc cat ggc cta gaa ttt gat	1104
209	Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp	
210	355 360 365	
212	tca aac aag gct atg gtt cct aaa cta gga act ggc ctt agt ttt gac	1152
213	Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp	
214	370 375 380	
216	agc aca ggt gcc att aca gta gga aac aaa aat aat gat aag cta act	1200
217	Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr	
218	385 390 395 400	
220	ttg tgg acc aca cca gct cca tct cct aac tgt aga cta aat gca gag	1248
221	Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu	
222	405 410 415	
224	aaa gat gct aaa ctc act ttg gtc tta aca aaa tgt ggc agt caa ata	1296
225	Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile	
226	420 425 430	
228	ctt gct aca gtt tca gtt ttg gct gtt aaa ggc agt ttg gct cca ata	1344
229	Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile	
230	435 440 445	
232	tct gga aca gtt caa agt gct cat ctt att ata aga ttt gac gaa aat	1392
233	Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn	
234	450 455 460	
236	gga gtg cta cta aac aat tcc ttc ctg gac cca gaa tat tgg aac ttt	1440
237	Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe	
238	465 470 475 480	
240	aga aat gga gat ctt act gaa ggc aca gcc tat aca aac gct gtt gga	1488
241	Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly	
242	485 490 495	
244	ttt atg cct aac cta tca gct tat cca aaa tct cac ggt aaa act gcc	1536
245	Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala	
246	500 505 510	
248	aaa agt aac att gtc agt caa gtt tac tta aac gga gac aaa act aaa	1584
249	Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys	
250	515 520 525	
252	cct gta aca cta acc att aca cta aac ggt aca cag gaa aca gga gac	1632
253	Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp	
254	530 535 540	
256	aca act cca agt gca tac tct atg tca ttt tca tgg gac tgg tct ggc	1680
257	Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly	
258	545 550 555 560	
260	cac aac tac att aat gaa ata ttt gcc acc tcg agt tac act ttt tca	1728
261	His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/914,151

DATE: 09/13/2001

TIME: 17:37:14

Input Set : A:\ES.txt

Output Set: N:\CRF3\09132001\I914151.raw

```

262          565          570          575
264   tac att gcc caa gaa cca tca gcc tcc gca tct gct tcc gcc cct gga   1776
265   Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly
266          580          585          590
268   tcc tac tcc atg gag cac ttc cgc tgg ggc aag ccg gtg taa   1818
269   Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val
270          595          600          605
273 <210> SEQ ID NO: 8
274 <211> LENGTH: 40
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence
278 <220> FEATURE:
279 <223> OTHER INFORMATION: synthetic DNA No.1037 used as sense primer for PCR
amplification of
280   DNA coding human MSH receptor residue 1-154.
282 <400> SEQUENCE: 8
283   gggaattcac catggctgtg cagggatccc agagaagact   40
286 <210> SEQ ID NO: 9
287 <211> LENGTH: 33
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: synthetic DNA No.1038 used as antisense primer for PCR
amplification
293   of DNA coding human MSH receptor residue 150-317.
295 <400> SEQUENCE: 9
296   gggaattcac caggagcatg tcagcacctc ctt   33
299 <210> SEQ ID NO: 10
300 <211> LENGTH: 27
301 <212> TYPE: DNA
302 <213> ORGANISM: Artificial Sequence
304 <220> FEATURE:
305 <223> OTHER INFORMATION: synthetic DNA No.1039 used as sense primer for PCR
amplification of
306   DNA coding human MSH receptor residue 150-317.
308 <400> SEQUENCE: 10
309   ctgcggtacc acagcatcgt gaccctg   27
311 <210> SEQ ID NO: 11
312 <211> LENGTH: 27
313 <212> TYPE: DNA
314 <213> ORGANISM: Artificial Sequence
316 <220> FEATURE:
317 <223> OTHER INFORMATION: synthetic DNA No.1040 used as antisense primer for PCR
amplification
318   of DNA coding human MSH receptor residue 1-154.
320 <400> SEQUENCE: 11
321   gctgtggtac cgcagtgcgt agaagat   27
324 <210> SEQ ID NO: 12
325 <211> LENGTH: 107
326 <212> TYPE: DNA
327 <213> ORGANISM: Artificial Sequence
329 <220> FEATURE:
330 <223> OTHER INFORMATION: synthetic DNA No.1075 used as sense primer for PCR

```

amplification of

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/914,151

DATE: 09/13/2001

TIME: 17:37:15

Input Set : A:\ES.txt

Output Set: N:\CRF3\09132001\I914151.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:33 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:36 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1